

## ERASING METHOD FOR NON-VOLATILE MEMORY

### CROSS-REFERENCE TO RELATED APPLICATION

5           This application claims the priority benefit of Taiwan application serial no. 91120417, filed September 9, 2002.

### BACKGROUND OF THE INVENTION

#### Field of Invention

10   [0001]       The present invention relates to a method of operating non-volatile memory. More particularly, the present invention relates to an erasing method for non-volatile memory.

#### Description of Related Art

[0002]       Non-volatile memory such as the electrically erasable programmable read  
15   only memory (EEPROM) is a device for holding data. Data stored inside the memory can be read or erased a multiple of times, and furthermore, the data is retained even after power is turned off. Hence, EEPROM has become one of the principal devices inside a personal computer or electrical equipment.

[0003]       The floating gate and the control gate of a typical EEPROM are fabricated  
20   using doped polysilicon. To program the memory, electrons injected into the floating gate will distribute evenly over the entire polysilicon floating gate layer. However, if the tunnel layer underneath the polysilicon floating gate layer contains defects, leakage current may flow and compromise the reliability of the device.

[0004] To prevent leakage current flowing inside an EEPROM, an electron-trapping layer is often used instead of a polysilicon floating gate layer. The electron-trapping layer is a layer of silicon nitride, for example. In general, a layer of silicon oxide is also formed both above and below the electron-trapping layer to form an oxide/nitride/oxide (ONO) composite stacked gate structure. An EEPROM having a stacked gate structure is often referred to as a nitride read-only-memory (nitride ROM). When voltages are applied to the control gate and the source/drain terminals of the device in a programming operation, hot electrons generated in the channel region close to the drain will inject into the electron-trapping layer. Since silicon nitride has the special property of capturing electrons, the injected electrons will concentrate within certain local areas of the electron-trapping layer instead of distributing evenly over the entire electron-trapping layer. Due to the localization of the electrons within the electron-trapping layer, sensitivity to defects in the tunnel oxide layer is reduced and hence leakage current is greatly minimized.

[0005] Another advantage of a nitride ROM is that the source/drain terminal on one side of the stacked gate may receive a higher voltage during a programming operation so that electrons are trapped inside the silicon nitride layer close to the source/drain terminal. Alternatively, the source/drain terminal on the other side of the stacked gate may receive a higher voltage so that electrons are trapped inside the silicon nitride layer close to that side of the source/drain terminal. Thus, by varying the voltage applied to the source/drain terminal on each side of the control gate, a single silicon nitride layer may be able to host two batches of electrons, a single batch of electrons or no electrons at all. In other words, a total of four states may be written into each memory cell of a nitride ROM. That means the nitride ROM is a two bits per cell type of non-volatile memory.

[0006] A conventional nitride ROM is programmed using channel hot electrons. Hot electrons injected from the drain side (or the source side) will penetrate through the tunnel oxide layer into the electron-trapping layer and are stored locally close to the area above the drain (or the source). Since the electron-trapping layer on the drain side (or the source side) has an excess of negative charges after the programming operation, threshold voltage ( $V_T$ ) of the memory cell will increase. These trapped electrons will remain inside the electron-trapping layer for a very long time (for example, up to 10 years at a temperature of 85°C) unless they are purposely erased in an erasing operation. To erase the electrons from the electron-trapping layer, holes on the drain side (source side) may penetrate the tunnel oxide layer into the electron-trapping layer in a band-to-band tunneling induced hot hole injection. Since the original negative electric charges inside the drain side (or source side) electron-trapping layer are neutralized by the holes after an erasing operation, threshold voltage ( $V_T$ ) of the memory cell will drop to re-constitute the erase state.

[0007] However, the quantity of holes injected from the drain side (or source side) into the electron-trapping layer is difficult to control if the band-to-band tunneling induced hot hole injection is utilized. Because of this, an excess of holes may be injected into the electron-trapping layer during erase and produce a so-called over-erase state in the memory device. When the degree of over-erasure is severe, reliability of the memory device may be compromised. The situation is particularly serious for a memory device having a small overall dimension. Thus, using band-to-band tunneling induced hot hole injection to carry out data erasure will ultimately limit the degree of miniaturization of memory devices.

## SUMMARY OF THE INVENTION

[0008] Accordingly, one object of the present invention is to provide an erasing method for non-volatile memory capable of preventing over-erasure, improving reliability and lowering operating current.

5 [0009] To achieve these and other advantages and in accordance with the purpose of the invention, as embodied and broadly described herein, the invention provides an erasing method for non-volatile memory. The non-volatile memory comprises a gate, a source, a drain, an electron-trapping layer and a substrate. The method includes applying a first voltage to the gate, a second voltage to the source, a third voltage to the drain and a fourth  
10 voltage to the substrate so that electrons are pulled out from the electron-trapping layer into the channel in the erasing operation. The difference between the first voltage applied to the control gate and the fourth voltage applied to the substrate is high enough to produce a negative gate F-N tunneling effect for erasing data within the non-volatile memory. Furthermore, the second voltage applied to the source, the third voltage applied to the drain  
15 and the fourth voltage applied to the substrate have identical value so that few holes are produced inside the substrate. Consequently, stress caused by the presence of holes is minimized.

[0010] The aforementioned erasing method is equally applicable to a non-volatile memory having two bits per cell storage capacity or a non-volatile memory having one bit  
20 per cell storage capacity.

[0011] This invention also provides a method of operating a non-volatile memory. The non-volatile memory includes a gate, a source, a drain, an electron-trapping layer and a substrate. The method includes programming the non-volatile memory to a programming

threshold voltage  $V_{t_p}$ . Thereafter, the non-volatile memory is set to an erased threshold voltage  $V_{t_e}$  through a negative gate F-N tunneling effect by applying a first voltage to the gate, a second voltage to the source, a third voltage to the drain and a fourth voltage to the substrate. The difference between first voltage applied to the gate and the fourth voltage applied to the substrate is high enough to produce a negative gate F-N tunneling effect for erasing data within the non-volatile memory. Furthermore, the second voltage applied to the source, the third voltage applied to the drain and the fourth voltage applied to the substrate have identical value so that few electric holes are produced inside the substrate and stress caused by the holes is restrained.

10 [0012] The aforementioned operating method is equally applicable to a non-volatile memory having two bits per cell storage capacity or a non-volatile memory having one bit per cell storage capacity.

[0013] Since the negative gate F-N tunneling effect raises the threshold voltage from the initial threshold voltage  $V_{t_i}$  (the initial state) to the erase threshold voltage  $V_{t_e}$  (the erase state), charges within the electron-trapping layer are already evenly distributed when the memory cells are in the erase state. This prevents residual holes existing in the memory cell of erased state. Furthermore, when negative gate F-N tunneling effect is used to erase data from the non-volatile memory, the erase threshold voltage  $V_{t_e}$  is a self-limiting level and hence effectively prevents any over-erasure.

20 [0014] The electron-trapping layer of the memory cell contains an excess of negative charges in both the programmed state and the erased state. Hence, electron transfer always occurs no matter if the memory cell changes from a programmed state to an erased state or vice versa. Consequently, reliability problems due to hole injection are prevented.

[0015] Moreover, the positive or negative gate F-N tunneling effect is used to initialize the memory cell after fabrication. Therefore, the erase threshold voltage  $V_{t_E}$  of the memory cell is greater than the initial threshold voltage  $V_{t_i}$ . Obviously, the initialization step may be skipped so that the erase threshold voltage  $V_{t_E}$  of the memory cell is identical to the  
5 initial threshold voltage  $V_{t_i}$ . In this case, the process of switching the memory cell from a programmed state to an erased state or vice versa still involves only electron transfer.

[0016] In addition, the erase method according to this invention may also be applied to a memory having a P-type gate or an N-type gate.

[0017] It is to be understood that both the foregoing general description and the  
10 following detailed description are exemplary, and are intended to provide further explanation of the invention as claimed.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings are included to provide a further understanding of the  
15 invention, and are incorporated in and constitute a part of this specification. The drawings illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention. In the drawings,

[0018] Figs. 1A to 1C are schematic cross-sectional diagrams showing the drain side bit programming, source side bit programming and data erasing operation inside a non-  
20 volatile memory cell according to this invention;

[0019] Figs. 2A to 2C are graphs showing the relationships between threshold voltage difference  $\Delta V_t(V_{t_P} - V_{t_E})$  and time for the drain side bit programming, source side

bit programming and data erasing operation  $\Delta Vt(Vt_E - Vt_p)$  inside a non-volatile memory cell according to this invention; and

[0020] Fig. 3 is a diagram showing the threshold voltages of a non-volatile memory according to this invention.

5

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0021] Reference will now be made in detail to the present preferred embodiments of the invention, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the description to refer to the same or like parts.

10

[0022] Figs. 1A to 1C are schematic cross-sectional diagrams showing the drain side bit programming, source side bit programming and data erasing operation inside a non-volatile memory cell according to this invention. Figs. 2A to 2C are graphs showing the relationships between threshold voltage difference  $\Delta Vt(Vt_p - Vt_E)$  and time for the drain side bit programming, source side bit programming and data erasing operation  $\Delta Vt(Vt_E - Vt_p)$

15

inside a non-volatile memory cell according to this invention. A non-volatile memory having a 2-bit per cell storage capacity is selected as an example to illustrate the embodiment of this invention. Refer to Figs. 1A to 1C and Figs. 2A to 2C for a better understanding of the operating modes of the non-volatile memory in this invention. In Figs. 2A to 2C, the icon ● (Bit-1) indicates a drain side bit and the icon ○ (Bit-2) indicates a source side bit. In this embodiment, programming is carried out utilizing channel hot electron injection effect and erasure is carried out utilizing negative gate F-N tunneling effect.

20

[0023] As shown in Fig. 1A, each memory cell in the non-volatile memory includes a substrate 100, a drain 102, a source 104, an electron-trapping layer 106, a gate 108, a tunneling oxide layer 110 and a blocking oxide layer 112. To program the drain side bit of a memory cell, a bias voltage  $V_{gp}$  of about 10V is applied to the gate 108, a bias voltage  $V_{dp}$  of about 4.5V is applied to the drain 102, a ground voltage of 0V is applied to the source 104 and a bias voltage  $V_{bp}$  of 0V is applied to the substrate 100. Under this bias condition, a large channel current (0.10 mA/memory cell to 1mA/memory cell) is generated. Electrons migrate from the source 104 towards the drain 102 and accelerate to create hot electrons by the intense channel electric field near the drain 102. Kinetic energy of the hot electrons is sufficiently high to overcome the energy barrier of the tunneling oxide layer 110. With the high positive bias voltage applied to the gate 108, the hot electrons at the drain 102 inject into an electron-trapping layer 106 and are stored within a localized region within the electron-trapping layer 106 above the drain 102. As negative charges accumulate in the electron-trapping layer 106 above the drain 102, threshold voltage of the memory cell begins to rise. In other words, the threshold voltage difference  $\Delta V_t(V_{t_p} - V_{t_e})$  of the drain side bit will increase with the programming time as shown in Fig. 2A.

[0024] Similarly, to program the source side bit of the memory cell as shown in Fig. 1B, a bias voltage  $V_{gp}$  of about 10V is applied to the gate 108, a bias voltage  $V_{sp}$  of about 4.5V is applied to the source 104, a ground voltage of 0V is applied to the drain 102 and a bias voltage  $V_{bp}$  of 0V is applied to the substrate 100. Under this bias condition, a large channel current is generated. Electrons migrate from the drain 102 towards the source 104 and accelerate to create hot electrons by the intense channel electric field near the source 104.



Kinetic energy of the hot electrons is sufficiently high to overcome the energy barrier of the tunneling oxide layer 110. With the high positive bias voltage applied to the gate 108, the hot electrons at the source 104 inject into an electron-trapping layer 106 and are stored within a localized region within the electron-trapping layer 106 above the source 104. As negative  
5 charges accumulate in the electron-trapping layer 106 above the source 104, threshold voltage of the memory cell begins to rise. In other words, the threshold voltage difference  $\Delta V_t(V_{t_p} - V_{t_e})$  of the source side bit will increase with the programming time as shown in Fig. 2B.

[0025] To carry out an erasing operation as shown in Fig. 1C, a bias voltage  $V_{ge}$  of  
10 between about -10V to -20V is applied to the gate 108, a bias voltage  $V_{de}$  and  $V_{se}$  of about 0V is applied to the drain 102 and the source 104 and a bias voltage  $V_{be}$  of about 0V is applied to the substrate 100. Thus, a potential difference is set up between the gate 108 and the substrate 100 creating a large electric field. Subsequently, negative gate F-N tunneling effect can be utilized to pull trapped electrons inside the electron-trapping layer 106 into the  
15 channel. During the erasing operation, a reduction of electrons inside the electron-trapping layer 106 leads to a drop in the threshold voltage of the memory cell. In other words, the threshold voltage difference  $\Delta V_t(V_{t_p} - V_{t_e})$  of the source side bit or the drain side bit will decrease with the programming time as shown in Fig. 2C.

[0026] In another erasing method according to this invention, a bias voltage  $V_{ge}$  of  
20 between about -10V to -20V is applied to the gate 108, a bias voltage  $V_{de}$  and  $V_{se}$  of between about 0V to 10V is applied to the drain 102 and the source 104 and a bias voltage  $V_{be}$  of between about 0V to 10V is applied to the substrate 100 (that is,  $V_{de} = V_{se} = V_{be}$ ).

Thus, a potential difference is set up between the control gate 108 and the substrate 100 creating a large electric field capable of pulling out trapped electrons inside the electron-trapping layer 106 into the channel. Using negative gate F-N tunneling effect to erase data has other advantages such as to operate the memory with a relatively small current.

5 [0027] In the method of erasing non-volatile memory according to this invention, the bias voltage  $V_{de}$  applied to the drain 102, the bias voltage  $V_{se}$  applied to the source 104 and the bias voltage  $V_{be}$  applied to the substrate 100 are all at the same value. Consequently, very few hot holes are created in the substrate 100 and hence stress due to hot hole effect is suppressed. Furthermore, the set up of a large electric field between the gate 108 and the  
10 substrate 100 (that is, a large electric field penetrating through the oxide/nitride/oxide (ONO) composite layer) facilitates the expulsion of electrons stored inside the electron-trapping layer (the silicon nitride layer) through the Frenkel-Pool emission effect. Thereafter, the electrons are pulled into the channel through the F-N tunneling effect. In addition, the aforementioned erasing method also provides a symmetrical bias to the source and the drain  
15 simultaneously (that is,  $V_{se} = V_{de}$ ) and thus rendering the simultaneous erasure of two bits of data previously registered inside a single memory cell possible. Obviously, the method can be applied to erase the data within a one bit per cell type of non-volatile memory.

[0028] Fig. 3 is a diagram showing the threshold voltages of a non-volatile memory according to this invention. As shown in Fig. 3, the threshold voltage of a memory cell  
20 immediately after fabrication is, for example, the initial threshold voltage  $V_{ti}$  (for example, 1.8V). An initialization step is next carried out changing the threshold voltage of the memory cell into an erase threshold voltage  $V_{te}$  (for example, 2.8V). The memory cell is initialized through the injection of electrons into the electron-trapping layer using a positive

or negative F-N tunneling effect so that the threshold voltage increases from the initial threshold voltage  $V_{ti}$  to the erase threshold voltage  $V_{tE}$ . Alternatively, the memory cell is initialized in two stages. First, electrons are injected into the electron-trapping layer using a positive or negative F-N tunneling effect or channel hot electron injection so that the threshold voltage increases from the initial threshold voltage  $V_{ti}$  to a programming threshold voltage  $V_{tp}$  (for example, 4.8V). Thereafter, an erase operation is carried out using a negative gate F-N tunneling effect to lower the threshold voltage of the memory cell from the programming threshold voltage  $V_{tp}$  to the erase threshold voltage  $V_{tE}$ .

[0029] To program the memory cell, channel hot electrons are injected into the electron-trapping layer above the drain side (or source side) terminal so that the threshold voltage of the memory cell increases further from the erase threshold voltage  $V_{tE}$  to the program threshold voltage  $V_{tp}$ . Thereafter, to erase data from the memory cell, the negative gate F-N tunneling effect is utilized to draw the electrons formerly injected into localized region of the electron-trapping layer out from the electron-trapping layer. Thus, the threshold voltage of the memory cell drops from the program threshold voltage  $V_{tp}$  back to the erase threshold voltage  $V_{tE}$ . In subsequent operation, the threshold voltage of the memory cell is raised to the program threshold voltage  $V_{tp}$  or lowered to the erase threshold voltage  $V_{tE}$  so that the memory cell is conditioned to a programmed state or the erased state.

[0030] In the process of raising the threshold voltage from the initial threshold voltage  $V_{ti}$  (the initial state) to the erase threshold voltage  $V_{tE}$  (the erase state) through the F-N tunneling effect, a layer of electric charges is evenly distributed within the electron-trapping layer. This prevents over-erasure and problems caused by the injection of excessive holes. Furthermore, the erase threshold voltage  $V_{tE}$  in the memory cell produced by

negative gate F-N tunnel effect is a self-limiting level and hence further prevents over-erasure.

[0031] Since the electron-trapping layer of the memory cell contains an excess of negative charges in both the programmed state and the erased state, electron transfer always occurs no matter if the memory cell changes from a programming state to an erasing state or vice versa. Consequently, reliability problems due to hole injection are prevented.

[0032] Moreover, the positive or negative gate F-N tunneling effect is used to initialize the memory cell after fabrication. Hence, the erase threshold voltage  $V_{tE}$  of the memory cell is greater than the initial threshold voltage  $V_{ti}$ . Obviously, the initialization step may be skipped so that the erase threshold voltage  $V_{tE}$  of the memory cell is identical to the initial threshold voltage  $V_{ti}$ . In this case, the process of switching the memory cell from a programmed state to an erase state or vice versa still involves only electron transfer.

[0033] In addition, the erase method according to this invention may also be applied to a memory having a P-type gate or an N-type gate.

[0034] It will be apparent to those skilled in the art that various modifications and variations can be made to the structure of the present invention without departing from the scope or spirit of the invention. In view of the foregoing, it is intended that the present invention cover modifications and variations of this invention provided they fall within the scope of the following claims and their equivalents.